DOCUMENT RESUME

ED 123 876 FL 006 683

AUTHOR Tarone, Elaine
TITLE A Discussion of the Dulay and Burt Studies. Working Papers on Bilingualism, No. 4.
INSTITUTION Ontario Inst. for Studies in Education, Toronto. Bilingual Education Project.
PUB DATE Oct 74
NOTE 15p.; For related documents, see FL 006 638, FL 006 681, FL 006 682, FL 006 684 and FL 006 685.

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS Bilingualism; Child Language; Error Patterns; Interference (Language Learning); *Language Development; *Learning Characteristics; *Learning Processes; *Learning Theories; Morphemes; *Psycholinguistics; *Second Language Learning; Statistical Analysis
IDENTIFIERS Burt (Marina); Dulay (Heidi)

ABSTRACT

Participants in a seminar series in second language acquisition held at Harvard University discussed three papers by Dulay and Burt ("Goofing: An Indicator of Children's Second Language Learning Strategies," "Should We Teach Children Syntax?", "Natural Sequences in Child Second Language Acquisition"), and developed several questions and issues regarding specific aspects of those papers. Among the issues raised are: (1) universal language processing strategies; (2) statistical discrepancies between the second and third studies; (3) interference in second language acquisition; (4) theoretical model of morpheme acquisition order; (5) morpheme scoring; (6) statistics and data analysis. Each question posed by the group is followed by a response from Dulay and Burt. (DB)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY Bilingual Education Project, OISE TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NATIONAL INSTITUTE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER."

ED123876

FL006683

A DISCUSSION OF THE DULAY AND BURT STUDIES

Elaine Tarone

with comments by Dulay and Burt

INTRODUCTION

Over the past three years, Heidi Dulay and Marina Burt have written four articles on second language acquisition, all of which are aimed at shaping a theory of second language acquisition as an alternative to the "habit formation" theory which underlies most of the contrastive analysis studies. Three of the articles are listed below:

(1) "Goofing: An Indicator of Children's Second Language Learning Strategies", Language Learning Vo. 22, No. 2 (1972) 235-252.

(2) "Should We Teach Children Syntax?", Language Learning Vo. 23, No. 2 (1973) 245-258.

(3) "Natural Sequences in Child Second Language Acquisition", Language Learning Vo. 24, No. 1 (1974) 37-53 and reprinted in Working Papers on Bilingualism No. 3 (1974) 44-67.

The participants* in a recent seminar series in second language acquisition held at Harvard University discussed these three papers in depth, and developed several questions and issues regarding specific aspects of those papers. The following list of questions and issues was sent to the authors for comment, and later discussed at the July 1974 meeting of the Linguistics Society of America at Amherst, Massachusetts.

Each question is followed by Dulay and Burt's response.

---

*Nancy Backman (Boston University), Jill and Peter de Villiers (Harvard), Kenji Hakuta (Harvard), Carmen Lanyon (Brandeis), Don Loritz (Boston University), Chuck Richards (Roxbury), John Schumann (Harvard), and Elaine Tarone (Boston University).

TARONE ET AL :

(1) In all three of your papers, you are attempting to demonstrate that when children learn a second language, they use the same "universal language processing strategies described in L1 research". (1973, p. 240) However, to make this comparison is to assume that we know what the strategies for L1 learning are. Peter and Jill de Villiers pointed out several times during the seminar that, in their own recent L1 research, they have found children doing all kinds of unexpected things in learning their first language--findings which have not been reported before, probably due to the limited number of subjects which have been studied in L1 research to date. The de Villiers' pointed out that it is probably premature to refer to "universal language processing strategies" used in L1 learning, and certainly premature to use such L1 learning strategies as a reference point for L2 research. Two possible solutions to this problem were considered: first, the Bilingual Syntax Measure (BSM) could be used with 30 month old children learning English as a first language, and then perhaps one could compare the results of L1 learners with L2 learners' results on the BSM. It was felt that only in this case could one truly compare the strategies of L1 learners with those of L2 learners. This sort of procedure might also provide some insight on ways in which the BSM might affect or change the strategies used by learners, since it does limit the obligatory context of features being studied in several ways. One would want to demonstrate that the sequence of morpheme acquisition is not just an artifact of the BSM testing situation; if one administered the test to L1 learners, for whom some spontaneous production data is being compiled, one might be able to determine whether the sequence of morphemes determined by the BSM is the same as that observed in spontaneous production of L1 learners in the research reported to date. Second, it might be helpful if you could be more specific about the "strategies for L1 learning" which you observe in L2 learning. That is, it might be more helpful to refer to specific processes like overgeneralization, strategy of communication, and so forth than to use the general term "L1 = L2".

DULAY
AND BURT : The "universal processing strategies" that were mentioned in the first two papers refer to the general statements about the L1 acquisition process that have been made repeatedly by L1 researchers. These have to do, first, with the finding that L1 acquisition is a process of "creative construction", not "simply a process of memorizing what has been heard" (Brown 1973: 98). "The most exciting conclusion from the study of child speech during the past ten years is that it is, in some respects, the same everywhere in the world...The most interesting generalizations from these studies concern the child's grammar or syntax." (Cazden and Brown, in press and Harvard ditto 1972: 1). "Children work out rules for the speech they hear, passing from levels of lesser to greater complexity, simply because the human species is programmed at a certain period in its life to operate in this fashion of linguistic input." (Brown 1973a: 105 and 1973b: 412) More specifically, L1 research has consistently found that children's reconstruction of the target language can be characterized by overgeneralizations, the use of pro-forms, the use of target language word order where word order is fixed (as in English or Samoan), the omission of functors, and in the earliest stage, the omission even of major constituents. These are the L1 findings we referred to when we posited similarities between the L1 and L2 acquisition processes. These similarities were strongly indicated by our comparison of errors made by child L2 learners of English with those reported in English L1 research (Dulay and Burt 1973 and 1974b).

Our L2 = L1 hypothesis was very specific and narrow in scope. As stated in the 1972 paper, it encompassed only syntactic error types--not the entire process of language acquisition. The 1973 paper in fact pointed out important differences between L1 and L2 acquisition, namely, it predicted and reported the finding that the acquisition order of 11 functors is radically different in L1 and L2 acquisition (pp. 254-255).

The major purpose of all three papers under discussion, but espe-

cially the first two, was to display the conflict between the habit formation and creative construction accounts of second language acquisition, and to attempt to resolve it. Although the creative construction view has become almost axiomatic for child language researchers today, the opposite view is still widely held in the second language teaching profession. Second language teachers usually do not have the time to become involved in ongoing basic research. The materials teachers work with are curriculum textbooks, not research reports, and most of these curriculum materials are based on the assumption that L2 learning proceeds by principles of habit formation, e.g. by imitation, repetition, reinforcement, immediate correction of any error, and transfer of first language behavior. We believe therefore, that it is of utmost importance for L2 researchers to gather data that would provide a sound empirical base to settle a real conflict in the field, as well as to provide a sound theoretical basis for second language teaching. Since L1 acquisition research provides much systematic evidence for the creative construction process, it would be foolish to ignore it in L2 research merely because there are certain obvious differences between the L1 and L2 learning processes. Of course, both L1 and L2 research will continue to improve on existing knowledge.

The use of the Bilingual Syntax Measure with L1 learners to make comparisons between L1 and L2 learning strategies seems like a good idea. We did not do so because the L1 findings in which we were interested had already been made through systematic and rigorous research. Further, though the BSM is an "assessment instrument", it does not use artificial tasks to elicit speech (e.g. the Berko "wug" test). The BSM elicits natural speech easily, much like one might use toys or picture books to stimulate conversation.

TARONE ET AL : (2) Tarone pointed out that there are important discrepancies between the results of Study 2 (1973) and Study 3 (1974). While the groups in Study 2 correlate highly with one another (in acquisition of 8 morphemes)

and the groups in Study 3 correlate highly with one another (in acquisition of 11 morphemes), the groups in Study 2 do not correlate well with the groups in Study 3. A Spearman rank order test produces the following correlations:

	Sacto Span (Study 2)	San Ysidro Span (2)	E Harlem Span (2)
Chinese (Study 3)	.57	.62	.60
Spanish (Study 3)	.77	.74	.74

The acquisition of morphemes by the Spanish speakers of the second study, when compared with the acquisition of morphemes by the Chinese speakers of the third study is not significant at .05 in similarity of sequence. The acquisition of morphemes by the Spanish speakers of the third study, however, is more similar to the results of the Spanish speakers of the second study--the similarity of sequence is just significant at .05.

These discrepancies between the two studies raise several questions. First, which order of morpheme acquisition is the correct one? That found in Study 2 or that found in Study 3? Second, why were the results different in the two studies? Was a different test instrument used in the two studies, or different experimenters? And do the results differ with different versions of the BSM, or different experimenters? Here again, it is important to know how the testing instrument affects the results. Third, it would appear that the Spanish speakers in the two studies were more similar to one another in their morpheme acquisition order than to the Chinese speakers when we measure across studies. Could this be considered evidence for the influence of L1 on the strategies used in acquiring L2, or interference?

DULAY AND BURT: We do not understand how the correlations presented above could have been obtained. Given that the Spearman rank order correlations between the Chinese and Spanish sequences in Study 3 were +.95 (Group Score Method) and +.96 (Group Means Method), it does not seem possible that significant differences could be obtained when comparing two similar sequences with any other sequences. Perhaps in the comparison of Studies 2 and 3, some mathematical error was introduced when the 11-functor sequence was reduced to an 8-functor sequence (which we presume was done to make the sequences of Studies 2 and 3 comparable).

Neither order of functors in the "correct" one, since there probably is no correct rank order or acquisition. Rather, it seems more likely that groups of functors are acquired together, and that it is groups of functors that are ordered rather than individual functors. (See Dulay and Burt in this issue--"New Perspectives" section.)

Given that the Spearman correlations you obtained are still in question, it might be advisable to approach the interference issue using clearer evidence. There is still much to be done in this area, in particular, the study of the effects of native language phonology on L2 morphology, as well as the reasons for the obvious greater amount of phonological interference in child L2 speech compared with the relatively insignificant amount of syntactic interference in the L2 speech of children (who are exposed to peers who natively speak the target language).

TARONE ET AL : (3) To varying degrees in the three papers, you seem to discount interference as a process in second language acquisition. What would you consider evidence for interference? What percentage of errors would have to be attributed to interference before it would become a valid process in second language acquisition? To a large degree, you seem to discount interference as a process in L2 acquisition, because you feel it would be evidence for a "habit formation" theory of language learning. However, the seminar participants felt that interference

could definitely be compatible with a creative approach to language learning. Corder describes language transfer as one of the hypotheses which learners might use in approaching the second language--that is, the hypothesis that the second language is like the first in certain ways.

DULAY AND BURT: The questions on the nature of acceptable evidence for interference cannot be answered in a theoretical vacuum. Statistics have meaning only within the context of a theoretical framework, e.g., how does one answer the question: How large is a large sample? The nature of acceptable confirming evidence for a hypothesis depends on the whole body of relevant known empirical facts and on the explanatory power of the resulting theory. (Salmon's (1973) article on "Confirmation" may be enlightening.)

We do not "discount interference as a process in L2 acquisition because /we/ feel it would be evidence for a habit formation theory of language learning." Rather, we discount habit formation as a process in L2 syntax acquisition because we found strong evidence against syntactic interference. Only 4.7% of 513 syntactic errors made by 179 children unambiguously reflected native language syntax, while 87.1% were similar to those described in L1 acquisition research. (Dulay and Burt, 1974b) This finding corroborates similar findings in a number of other studies. There is also evidence against positive transfer (Richards 1971, Wolfe 1967, Hernandez 1972).

If your notion of language transfer is the one subscribed to by cognitive psychologists, that is, that past experience plays a significant role in new learning experiences, then we would entirely agree.

TARONE ET AL : (4) Given that one of the orders of acquisition of morphemes isolated in your studies is more plausible than the other, what sort of theoretical construct would you provide to explain why that particular order of acquisition is used? Is any work being done on this kind of theoretical construct? Superficially, the order doesn't seem to accord with phonological distinctiveness, syllabic-ness, frequency, semantic or

syntactic complexity. An invariant order of acquisition without an explanation for the order seems somewhat incomplete.

DULAY AND BURT: This would require a lengthy discussion. Please see our paper in this issue where we touch on some of these important topics.

TARONE ET AL :(5) There were a number of questions on technical or procedural points relating to your studies using the Bilingual Syntax Measure. A number of these questions arose from researchers currently involved in studies of L2 acquisition, many of which involve the counting of morphemes. Subjects--How much ESL instruction had the subjects had in a formal context? How similar were they to one another?

DULAY AND BURT: We assume you are referring to the subjects in Study 3. All subjects received some formal ESL instruction, varying from 2-3 hours per week to 2 hours per day. As to their similarity, what variables do you have in mind?

TARONE ET AL :Scoring

(a) In any study that measures morphemes there are problems involved in scoring higher-order errors, such as word-order errors or semantic errors. If a child inverts or modifies word order from the target language, does the study just count morphemes and not take word order errors into account? In Study 1 (Table 1) (1973), some examples of changes in word order by L2 learners are given. Are these changes in word order scored wrong on a morpheme-counting measure? Also, how do you score responses of the child which show that he/she clearly doesn't understand the question? (e.g., "Which one is he?" "She's the mother".)

DULAY AND BURT: In scoring morphemes, we introduced a refinement on Brown's method, that is, misformed functors (which include misordering of a given functor, or semantic errors) were given half the value of a correctly supplied functor. Word order (and other higher level) errors do not affect the morphemes (e.g. the dog the food eat, where the article is given full value. Word order errors can be tallied separately. Responses which show that the child did not under-

stand the question were discarded from the analysis.

TARONE ET AL :(b) The tabulation of the number of morphemes produced correctly in obligatory context seems to obscure some important data--the types of errors made when morphemes are not produced correctly. For example, the number of errors made with articles is tabulated, but not the type of error made with articles. In an ongoing study by D. Loritz with adult Chinese learners of ESL, there is a distinct overgeneralization of "the" articles (out of 507 missupplied articles, 453 entailed substituting "the" inappropriately), while all indications are that L1 learners do the opposite, overgeneralizing indefinite articles. The morpheme-counting measure cannot really differentiate (in its present form) the different types of errors evidencing different strategies which might be present in learners from different language backgrounds.

DULAY AND BURT: Yes.

TARONE ET AL :(c) In all second language acquisition studies, there is the problem of what to do with "performance clutter"--that is, the utterances of learners Evelyn Hatch calls "data gatherers"--such as "He who the fat skinny". In your study, you simply threw those utterances out, since, in order to count the number of morphemes produced correctly in obligatory context, it is necessary to be able to determine what structures the child is offering. However, this might be disposing of valuable data in L2 acquisition. Those utterances being thrown out are those which are too far removed from the target language structure for the experimentor to interpret; might this not be biasing the measure toward finding only those errors which are close to the L2 structure? Might those "performance clutter" errors contain valuable information about learners' strategies to communicate meaning? As such, it was felt that researchers in L2 acquisition ought to at least record the number of such utterances which were not tabulated in a study, and ought to make that data available in some form to other researchers.

DULAY AND BURT: We agree. Incidentally, each child's protocol in the national BSM Field Test included the number of such incomprehensible utterances.

TARONE ET AL :(d) Were the scores of individual children tabulated separately and compared to one another? And if so, was the sequence of acquisition of an individual child predictable from the sequence of morphemes displayed in Study 3, for example? In Study 3, Table 3-5 (1973), there appears to be a wide variability in acquisition--individual children seem to have been very different.

DULAY AND BURT: No. This task still remains to be done.

TARONE ET AL :(e) Since the scoring of your studies is different in certain respects from the scoring of L1 data in the Brown and de Villiers studies, does this affect the comparability of your findings with theirs? For example, in your studies you weight the scores from 0, 1, to 2. In the L1 studies, errors are either 0 or 1.

DULAY AND BURT: This cannot be answered without rescoring the data. We made the refinement because we believed it would add significantly to the accuracy of the L2 data.

TARONE ET AL :(f) Why didn't you use all 14 morphemes from the Brown study? Was there a semantic factor, or a frequency factor?

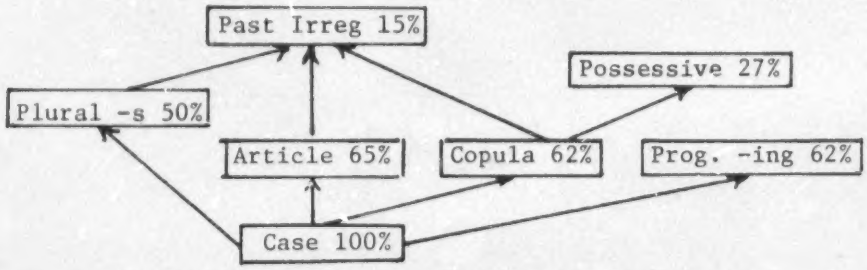
DULAY AND BURT: We were able to regularly elicit only the functors reported in the studies.

TARONE ET AL :(g) How did you score articles? Were incorrect suppliances (the/a) scored, and if so, how? Also, how was an utterance "he is" scored in a contractible context? 0, 1 or 2? A need for careful definitions of obligatory contexts and scoring procedures was expressed by some researchers interested in comparing their own morpheme studies with yours, de Villiers' and Brown's.

DULAY
AND BURT: Where the distinction between the and a was clear, we scored them according to the metric described. We included "he is..." in the tallying of contractible copula, but not, for example, "there he is".

TARONE
ET AL :Statistics and Data Analysis

(a) Kenji Hakuta raised some questions about the "Tree Method to Determine Acquisition Hierarchies"--a handout at the 1974 TESOL Convention which accompanied the presentation of the results of your Study 3.



This method of presenting data in L2 acquisition studies was felt to hold a great deal of promise. However, Hakuta pointed out that if Case is omitted from this tree, all you're saying is that the Past irregular comes in late, and the Possessive comes in after the Copula. Is it meaningful to consider case in a morpheme study? P. de Villiers said that he would soon have data on case in L1 acquisition, but that at present it seems that case doesn't show the same invariance in order that the morphemes show in L1 learning. That is, in L2 learning, pronouns seem to be quite variable--good learners may get them late, and poor learners may get them early. Given the problems involved in generalizing from L1 studies to L2 studies, the de Villiers findings may raise some problems with the inclusion of case in an L2 study of morpheme acquisition.

DULAY
AND BURT: The purpose of our studies of acquisition sequence was not to compare the acquisition of morphemes vs. non-morphemes. We would have included higher level structures had the methodology for handling such

structures been available. Case was included in our study because it was scorable in the same way the other morphemes were.

TARONE ET AL : (b) Why was a rank-order measure used rather than a product-moment measure of correlation? Since percentage measures are available, the product-moment measure could be used and would be more accurate. The Spearman Rank Order formula produces a "group means" linear correlation of .96, but the product-moment formula produces a correlation of .87--still very significant, but not as startling as .96. The Spearman formula may tend to obscure some divergencies in the data of the 1974 paper.

DULAY AND BURT: A rank order measure was used because it had been used in the de Villiers' cross-sectional study of acquisition sequence. Further, our statistical consultants (from Harvard University and the Far West Laboratory for Educational Research and Development, San Francisco) did not advise use of the product-moment measure.

TARONE ET AL : (c) Presenting the data in terms of correlations may also tell only part of the story. That is, since x = y and x = y + 10 correlate with r = 1.0, correlation tells nothing about constants. In your 1974 paper, Loritz pointed out that the difference in mean scores averages to 23%, that is, the Chinese did 23% worse than the Spanish in inflecting English. Loritz does not believe that we can yet reject the hypothesis that this is at least partly due to the fact the Chinese, unlike Spanish and English, is an utterly uninflected language.

DULAY AND BURT: Correlations perform a specific function in the acquisition order studies we have been discussing. That is, they are able to show the relation between rank orders regardless of the mean level of proficiency of the samples compared. The Chinese children in Study 3 were indeed less proficient on the average than the Spanish children. (The Chinese children had lived in the U. S. less time than the Spanish children.)

The fact that the rank orders correlated as highly as they did makes the findings about the sequences even stronger. A comparison of mean scores might be useful for studies with an entirely different purpose.

REFERENCES

Brown, R. 1973. A First Language. Cambridge, Mass.: Harvard University Press.

Brown, R. 1973. Development of the first language in the human species. American Psychologist, February. 97-102.

Burt, M., H. Dulay and E. Hernandez. 1973. Bilingual Syntax Measure (Restricted Edition). New York: Harcourt Brace Jovanovich.

Cazden, C.B. and R. Brown. in press. The early development of the mother tongue. In Lenneberg, Eric and Elizabeth (eds.), Foundations of Language Development: A Multidisciplinary Approach. UNESCO. Also ditto, Harvard University, 1972.

De Villiers, J. and P. de Villiers. 1973. A cross-sectional study of the acquisition of grammatical morphemes in child speech. Journal of Psycholinguistic Research 2.

Dulay, H. and M. Burt. 1972. Goofing: an indicator of children's second language learning strategies. Language Learning 22 (2). 235-252.

Dulay, H. and M. Burt. 1973. Should we teach children syntax? Language Learning 23(2). 245-258. Reprinted in Light, R.L. and M. Gutierrez (eds.), Handbook for ESL Teachers, New York State Department of Education, in press.

Dulay, H. and M. Burt. 1974a. Natural sequences in child second language acquisition. Language Learning 24(1). 37-53. Reprinted in Working Papers on Bilingualism No. 3, 44-67, Ontario Institute for Studies in Education.

Dulay, H. and M. Burt. 1974b. Errors and strategies in child second language acquisition. TESOL Quarterly 8(2). 129-136.

Dulay, H. and M. Burt. 1974. A new perspective on the creative construction process in child second language acquisition. Working Papers on Bilingualism, No. 4, 71-98, Ontario Institute for Studies in Education.

Hernandez-Chavez, E. 1972. Early code separation in the acquisition of English by Spanish-speaking children. Paper presented at the Stanford Child Language Research Forum, Stanford University, Stanford, California.

Richards, J. 1971. Error analysis and second language strategies. Language Sciences 17.

Salmon, W.G. 1973. Confirmation. Scientific American February: 75-83.

Wolfe, D.L. 1967. Some theoretical aspects of language learning and language teaching. Language Learning 17.